Sediment Budgets (Proceedings of the Porto Alegre Symposium, December 1988). IAHS Publ. no. 174, 1988.

The reliability of rating curve estimates of suspended sediment yield: some further comments

D. E. WALLING & B. W. WEBB
Department of Geography, University of Exeter, Exeter EX4 4RJ, UK

Abstract A number of recent papers have suggested that the underestimation bias inherent in the use of log-transformed regression to derive suspended sediment rating relationships is a major cause of the error commonly associated with rating curve estimates of suspended sediment yield, and that this bias can be largely removed by applying simple correction factors. However, there is conflicting evidence as to the extent to which bias-corrected rating estimates provide reliable values of sediment yield. In order to investigate this uncertainty further, the authors have assembled continuous records of suspended sediment concentration and discharge for three rivers in Devon, UK. These records have been used to calculate actual loads and to synthesize a variety of replicate data sets, representing different manual sampling strategies, for rating curve derivation. The resulting load estimates obtained using rating curves and bias-corrected rating curves have been compared with the actual loads. The results indicate that bias-correction procedures do not provide accurate estimates for these rivers and that other sources of error associated with rating curves are more important in producing inaccurate estimates.

La validité des données relatives aux charges solides en suspension obtenues en utilisant les courbes de tarage: quelques commentaires supplémentaires

Résumé Plusieurs articles récents ont prétendu que la tendance à la sous-estimation inhérente à l'emploi de la régression des logarithmes pour dériver les courbes de tarage est une cause majeure des erreurs associées fréquemment aux données relatives aux charges solides en suspension obtenues à partir des courbes de tarage, et que cette tendance peut être corrigée par l'application de facteurs de correction simples. Cependant il n'est pas évident que les courbes de tarage corrigées de cette tendance puissent fournir des estimations exactes. Afin d'examiner cette incertitude plus à fond, les auteurs ont assemblé des données continues des concentrations des sédiments en suspension et des écoulements pour trois rivières du

Devon, Royaume Uni. Ces données ont été employées à calculer les valeurs varies de transports solides en suspension et à faire la synthèse de diverses séries de données duplicatives correspondant à différentes stratégies d'échantillonnage manuel en vue de déterminer la courbe d'étalonnage. Les estimations de transports solides en suspension obtenues à partir de courbes de calibration sans correction et les courbes corrigées pour la tendance à la sous-estimation ont été comparées avec les valeurs véritables. Les résultats indiquent que les opérations effectuées pour corriger la tendance à la sous-estimation ne fournissent pas les estimations exactes pour ces rivières et que d'autres sources d'erreur associées aux courbes de tarage ont une plus forte influence sur les erreurs d'estimation.

INTRODUCTION

The study of Colby (1956) provides an excellent review of the many problems involved in any attempt to develop a quantitative relationship between suspended sediment discharge and water discharge and therefore in using sediment rating curves to estimate suspended sediment yields. His work has been followed by many other investigations which have clearly demonstrated that rating curves are unlikely to provide accurate estimates of sediment yield (cf. Farr & Clarke, 1984; Olive *et al.*, 1980; Walling & Webb, 1981). For example, Walling & Webb (1981) present the results of a study of the suspended sediment load of the River Creedy in Devon, UK, which indicated that rating curves could provide load estimates for a seven-year period which were only 20% of the actual or true load. These authors also refer to two different attempts to estimate the suspended sediment load of the River Cleddau in South Island New Zealand, based on rating curves (Griffiths, 1979; Adams, 1980), which produced values differing by nearly two orders of magnitude. More recently, Ferguson (1984, 1986a, 1987) has highlighted an underestimation bias inherent in the use of log-transformed regression to derive rating relationships. He has suggested that this bias represents a major cause of the error associated with rating curve estimates and that it can be largely removed by applying a simple correction factor based on the standard error of the estimate of the logarithmic regression. For common logarithms, the correction factor required to calculate the unbiased estimator of the true load (L_{cr}) is derived as:

$$CF_1 = \exp(2.651\ s^2) \qquad (1)$$

where s is the standard error of the estimate of the rating curve in $\log_{10}$ units. Koch & Smillie (1986) have provided an alternative non-parametric function for bias correction based on the smearing estimate of Duan (1983). This correction factor is calculated as:

$$CF_2 = (1/n)\ \sum_{i=1}^{n} 10^{e_i} \qquad (2)$$

$$e_i = \log(C_i) - \log(Ce_i) \qquad (3)$$

where $\log(C_i)$ is the log of concentration observation i, $\log(Ce_i)$ is the estimated log value of concentration for the same observation derived from the regression using the appropriate value of discharge, and n is the number of observations. Hansen & Bray (1987) observe that both correction factors approach unity as the scatter about the best-fit line decreases and the standard error of the estimate and the correlation coefficient approach zero and unity respectively. In addition, they note that the values of both factors will be the same if the residuals are normally distributed about the best-fit line. In each case, the unbiased estimator of the true load is obtained by multiplying the original rating curve estimate by the correction factor.

In his analysis of the merits of applying the bias correction factor CF_1, Ferguson (1986a, 1987) tested the accuracy of bias corrected and standard rating curve estimates of sediment yield against the actual load for the streams involved. The results obtained were impressive. The bias correction factor consistently reduced the underestimation associated with the standard rating curve estimates and the resultant estimates were within 10% of the true values. Similar findings were reported by Hansen & Bray (1987) who reported that use of the smearing estimate correction factor (CF_2) produced rating curve estimates of suspended sediment yield for the Kennebecasis River in New Brunswick, Canada, which were within about 10% of the loads published by the Sediment Survey of Canada. These findings have encouraged others to employ the bias correction procedure and to assume that the resultant load estimates provide reliable values of sediment yield which can be employed in establishing sediment budgets (e.g. Stott *et al.*, 1986).

Others, however, have been less convinced of the capacity of bias corrected rating estimates to provide reliable values of sediment yield. For example, Koch & Smillie (1986) report an analysis of sediment yield estimates for the Yampa and Little Snake Rivers in northwest Colorado, USA, which indicated that both bias correction factors produced sediment yield estimates which significantly overestimated the actual mean annual sediment yields and which were generally less accurate than those provided by the standard rating curve. Ashmore (1986) reports a comparison of bias corrected (CF_1) rating curve estimates of mean annual sediment yields and the equivalent loads published by the Sediment Survey of Canada for a number of rivers in Saskatchewan which indicated errors in the former estimates within the range ±50%. In this case, as indeed for the study involving the Kennebacasis River noted above, the published data cannot be viewed as providing an absolute standard against which to assess the accuracy of the rating curve estimates, but the accuracy of these estimates is clearly placed in doubt. Another Canadian study by Kellerhals Engineering Services (1985) of six rivers in British Columbia casts further doubt on the merits of applying the bias correction factor (CF_1) to rating curve estimates of annual sediment yield. Sixty-six stations years of data were analysed and the annual sediment loads published by the sediment survey of Canada were used as the standard. In very few cases did the bias-corrected rating curve estimate provide a closer approximation than the standard rating curve estimate. The latter were nearly always underestimates, whilst in the majority of cases the former were overestimates. For example, in the case of Grave Creek, rating curves were derived for eight individual years. The standard rating curve estimates underestimated the published annual loads by between −17 and −62%, with an average of −39% but the bias-corrected estimates overestimated by between +6 and +268%, with an average of +91%.

Uncertainty therefore exists both as to the merits of applying a bias correction factor to rating curve load estimates and to the extent to which bias represents a major cause of the inaccuracy associated with rating curve estimates. Some of the evidence outlined above suggests that other sources of error (cf. Walling, 1977a, b) are more significant and that correction for bias is in itself unlikely to produce reliable estimates of sediment yield. In order to investigate this uncertainty further, the authors have assembled continuous records of sediment concentration and water discharge for three rivers in Devon, UK (Table 1). The rivers provide a range of basin sizes and the continuous records of sediment concentration obtained from calibrated turbidity monitors cover substantial periods of up to 10 years. These records have been used to calculate accurate "actual" loads with which rating curve estimates can be compared and to synthesize a variety of replicate data sets, representing different manual sampling strategies, for rating curve derivation.

Table 1 *Rivers involved in the analysis*

River	*Gauging station*	*Area*	*Period of sediment record*
Dart	*Bickleigh*	*46*	*1975-1985*
Creedy	*Cowley*	*262*	*1972-1980*
Exe	*Thorverton*	*601*	*1978-1980*

DATA ANALYSIS

The continuous records of river discharge and suspended sediment concentration for the three rivers listed in Table 1 have been digitized at hourly intervals to provide continuous series. Values of "actual" annual load have been obtained by integrating the discharge and concentration series at hourly increments. Two sampling strategies have been employed to synthesize typical representative data sets for rating curve construction. The first involves sampling at regular weekly intervals whilst the second supplements a programme of regular weekly sampling with additional samples collected during periods of flood discharge. These periods have been identified by specified flow thresholds and random numbers have been used to select sampling times during the occasions when these thresholds are exceeded. Table 2 provides further details of the number of samples collected during flood events and the thresholds employed. Fifty replicate data sets have been generated for both sampling strategies to permit evaluation of both the accuracy and precision of the rating curve estimates (cf. Walling & Webb,

Table 2 *Sampling strategies used for generating flood period data sets*

River	*Flow class 1 ($m^3 s^{-1}$)*	*Flow class 2 ($m^3 s^{-1}$)*	*No. in class 1*	*No. in class 2*
Dart	*5-10*	*>10*	*200*	*150*
Creedy	*15-30*	*>30*	*160*	*120*
Exe	*30-60*	*>60*	*40*	*30*

1981). In the case of regular weekly sampling, these replicates were obtained by undertaking the sampling at different times during the week. For the flood event sampling, random numbers were used to select different sampling times for each replicate during the periods when the flow thresholds were exceeded. Comparison of "actual" and rating curve estimates of sediment load have been limited to annual and multi-annual values since rating curves are rarely used to estimate shorter-term loads. Analysis has also been restricted to the derivation of rating curves for the entire period of record, rather than for individual years, since this is common practice when only a relatively small number of samples are available.

All rating relationships were established using least squares linear regression on the log transformed data sets. Loads were estimated for the rating relationships by applying the hourly discharge series to the relationship to estimate the suspended sediment concentration and by summing the resultant estimates of hourly load. Bias correction was applied to these aggregate loads by using the two correction factors (CF_1 and CF_2) introduced above.

RESULTS

The results obtained for the three rivers in using rating relationships based on the 50 replicate data sets to estimate annual sediment yields and the total load for the period of record are presented in Table 3 for the regular sampling strategy and in Table 4 for the regular plus flood event sampling strategy. Table 3 indicates that use of a simple uncorrected logarithmic rating relationship fitted to data obtained from a programme of regular weekly sampling produces serious underestimation of annual and total sediment yields for all three rivers. In the case of the River Dart, the means of the 50 replicate rating curve estimates are typically only about 5% of the actual annual loads. Although the degree of underestimation appears to decrease with increasing basin size, and is therefore slightly less for the other two rivers, the mean values of total load for the period of record estimated using the rating relationship are only 19.5% and 7.3% of the actual loads of the River Creedy and the River Exe respectively. There is also considerable imprecision associated with the load estimates. The coefficient of variation of the 50 replicate estimates of annual and total load is typically of the order of 15% in the case of the Rivers Dart and Creedy and even higher (*c.* 25%) for the River Exe. Table 3 indicates reasonable consistency in the degree of

***Table 3** The mean and coefficient of variation of the suspended sediment loads estimated using the rating curves based on the 50 replicate data sets representing regular weekly sampling*

Period	*Actual load*	*Simple rating*	*CF_1 correction*	*CF_2 correction*
	(t)	*$\bar{x}(t)$ cv(%)*	*$\bar{x}(t)$ cv(%)*	*$\bar{x}(t)$ cv(%)*
River Dart at Bickleigh				
1975-1985	*24 499*	*897 (17.0)*	*2 308 (27.5)*	*5 072 (23.9)*
1975-1976	*1 072*	*33 (13.4)*	*84 (24.1)*	*186 (21.1)*
1976-1977	*3 779*	*130 (19.2)*	*336 (29.6)*	*737 (25.8)*
1977-1978	*1 872*	*75 (15.7)*	*193 (26.3)*	*425 (22.9)*
1978-1979	*1 476*	*96 (19.1)*	*246 (29.5)*	*540 (25.7)*
1979-1980	*2 475*	*98 (16.5)*	*251 (27.1)*	*551 (23.5)*
1980-1981	*2 684*	*96 (16.4)*	*248 (26.9)*	*545 (23.4)*
1981-1982	*4 451*	*143 (17.5)*	*367 (28.1)*	*806 (24.4)*
1982-1983	*2 672*	*95 (15.9)*	*245 (26.5)*	*538 (23.0)*
1983-1984	*3 046*	*86 (16.6)*	*221 (27.2)*	*487 (23.6)*
1984-1985	*972*	*45 (14.5)*	*117 (25.2)*	*257 (22.0)*
River Creedy at Cowley				
1972-1980	*82 863*	*16 125 (15.2)*	*41 257 (14.5)*	*49 769 (29.1)*
1972-1973	*7 482*	*1 070 (14.5)*	*2 739 (13.9)*	*3 303 (28.8)*
1973-1974	*20 619*	*3 395 (15.9)*	*8 685 (15.2)*	*10 479 (29.5)*
1974-1975	*10 547*	*1 603 (13.9)*	*4 102 (13.3)*	*4 947 (28.5)*
1975-1976	*1 941*	*163 (10.5)*	*418 (10.3)*	*503 (26.9)*
1976-1977	*16 234*	*3 302 (14.6)*	*8 448 (14.0)*	*10 190 (28.8)*
1977-1978	*10 214*	*3 198 (16.4)*	*8 182 (15.6)*	*9 874 (29.8)*
1978-1979	*4 717*	*1 108 (12.7)*	*2 835 (12.2)*	*3 417 (27.9)*
1979-1980	*11 109*	*2 286 (15.9)*	*5 847 (15.2)*	*7 055 (29.5)*
River Exe at Thorverton				
1978-1980	*41 402*	*3 010 (25.4)*	*10 472 (36.0)*	*17 648 (31.0)*

underestimation between the individual years. In the case of the River Dart the means of the replicate estimates of annual load indicate loads ranging from 2.8–6.5% of the actual load over the 10-year period. There is, however, greater inter-year variability in the degree of underestimation evident for the River Creedy, with equivalent values of 8.4–31.3%. The effects of including additional flood event samples in the data sets used to derive the rating relationships can be examined by considering the results presented in Table 4. In the case of the Rivers Dart and Creedy, there is some limited improvement in both the accuracy and precision of the load estimates, but this is more marked for the River Exe, where the degree of underestimation associated with the mean of the replicate loads for the period of record decreases from –92.7% to –81% and the coefficient of variation of the same loads reduces from 25.4% to 18.8%. The degree of interannual variability in the level of accuracy of the load estimates is also substantially reduced by the addition of the flood event samples. Overall, however, the results obtained for both sets of rating relationships indicate that the standard rating curve technique is unlikely to provide meaningful estimates of annual and total sediment yield for these three rivers.

The effects of applying the two bias correction factors to the above load estimates can be seen in Tables 3 and 4. In the case of the rating relationships based on a programme of weekly sampling (Table 3), the degree

Table 4 *The mean and coefficient of variation of the suspended sediment loads estimated using the rating curves based on the 50 replicate data sets representing regular weekly plus flood period sampling*

Period	*Actual load*	*Simple rating*		*CF_1 correction*		*CF_2 correction*	
	(t)	*$\bar{x}(t)$*	*cv(%)*	*$\bar{x}(t)$*	*cv(%)*	*$\bar{x}(t)$*	*cv(%)*
River Dart at Bickleigh							
1975-1985	*24 499*	*1 570*	*(12.9)*	*5 277*	*(19.5)*	*8 366*	*(15.4)*
1975-1976	*1 072*	*49*	*(10.3)*	*166*	*(17.1)*	*263*	*(13.6)*
1976-1977	*3 779*	*250*	*(14.5)*	*840*	*(21.0)*	*1 330*	*(16.6)*
1977-1978	*1 872*	*124*	*(11.8)*	*418*	*(18.5)*	*662*	*(14.6)*
1978-1979	*1 476*	*182*	*(14.4)*	*613*	*(20.9)*	*971*	*(16.5)*
1979-1980	*2 475*	*167*	*(12.5)*	*561*	*(19.1)*	*890*	*(15.1)*
1980-1981	*2 684*	*164*	*(12.3)*	*551*	*(18.9)*	*874*	*(14.9)*
1981-1982	*4 451*	*255*	*(13.3)*	*859*	*(19.8)*	*1 361*	*(15.7)*
1982-1983	*2 672*	*159*	*(12.0)*	*533*	*(18.6)*	*845*	*(14.7)*
1983-1984	*3 046*	*148*	*(12.6)*	*497*	*(19.1)*	*788*	*(15.1)*
1984-1985	*972*	*72*	*(11.0)*	*240*	*(17.7)*	*381*	*(14.1)*
River Creedy at Cowley							
1972-1980	*82 863*	*23 936*	*(13.5)*	*61 766*	*(12.4)*	*68 018*	*(23.3)*
1972-1973	*7 482*	*1 556*	*(13.0)*	*4 014*	*(11.9)*	*4 420*	*(23.0)*
1973-1974	*20 619*	*5 162*	*(14.2)*	*13 319*	*(13.0)*	*14 669*	*(23.7)*
1974-1975	*10 547*	*2 274*	*(12.3)*	*5 869*	*(11.3)*	*6 462*	*(22.7)*
1975-1976	*1 941*	*206*	*(9.3)*	*533*	*(8.8)*	*587*	*(21.3)*
1976-1977	*16 234*	*4 793*	*(12.9)*	*12 370*	*(11.9)*	*13 620*	*(23.0)*
1977-1978	*10 214*	*4 949*	*(14.7)*	*12 769*	*(13.5)*	*14 064*	*(23.9)*
1978-1979	*4 717*	*1 506*	*(11.1)*	*3 888*	*(10.3)*	*4 279*	*(22.1)*
1979-1980	*11 109*	*3 490*	*(14.4)*	*9 005*	*(13.2)*	*9 918*	*(23.8)*
River Exe at Thorverton							
1978-1980	*41 402*	*7 886*	*(18.8)*	*48 148*	*(17.5)*	*35 775*	*(20.6)*

of underestimation of the annual and total loads is reduced by applying the correction factor CF_1, although the rating curve estimates of annual load are typically still only about 10% of the actual load for the River Dart, 25% for the River Exe and 50% for the River Creedy. This correction factor results in a small improvement in the precision of the load estimates for the River Creedy, but the precision is markedly worse for the Rivers Dart and Exe, where the coefficients of variation of the replicate estimates of annual load are typically about 26% and 36% respectively. The interannual variability in the levels of accuracy and precision associated with the load estimate is also increased by application of this correlation factor. Use of the smearing estimate correction (CF_2) on the load estimates provided by the rating relationships developed from regular weekly sampling produces similar results to the correction factor CF_1, although the corrected load estimates are generally better with respect to accuracy and sometimes worse in terms of precision. For the River Dart, however, the mean of the replicate estimates of annual load is still typically only about 20% of the actual load and the coefficient of variation of the replicate estimates is commonly about 25%.

Use of the two bias correction factors with the rating relationships produced using the data sets involving both regular weekly samples and flood event samples (Table 4) produces some improvement in accuracy and precision, and the mean of the replicate total load estimates for the River Exe obtained using CF_1 overestimates the actual load by 16.3%. However, the results are still far from encouraging. For the River Dart, the means of the replicate annual load estimates obtained using both correction factors are still typically only 20–30% of the actual values and in general there is an increase in the interannual variation in the degree of accuracy of the load estimates. For the River Creedy, application of CF_1 produces mean values of the replicate load estimates for individual years which range between 27.5% and 125% of the actual values, whereas with CF_2 the equivalent values are 30.2% and 137.7%.

DISCUSSION

The results presented in Table 3 and 4 indicate that neither standard logarithmic rating relationships nor logarithmic relationships corrected for bias using the correction factors CF_1 and CF_2 are able to provide reliable

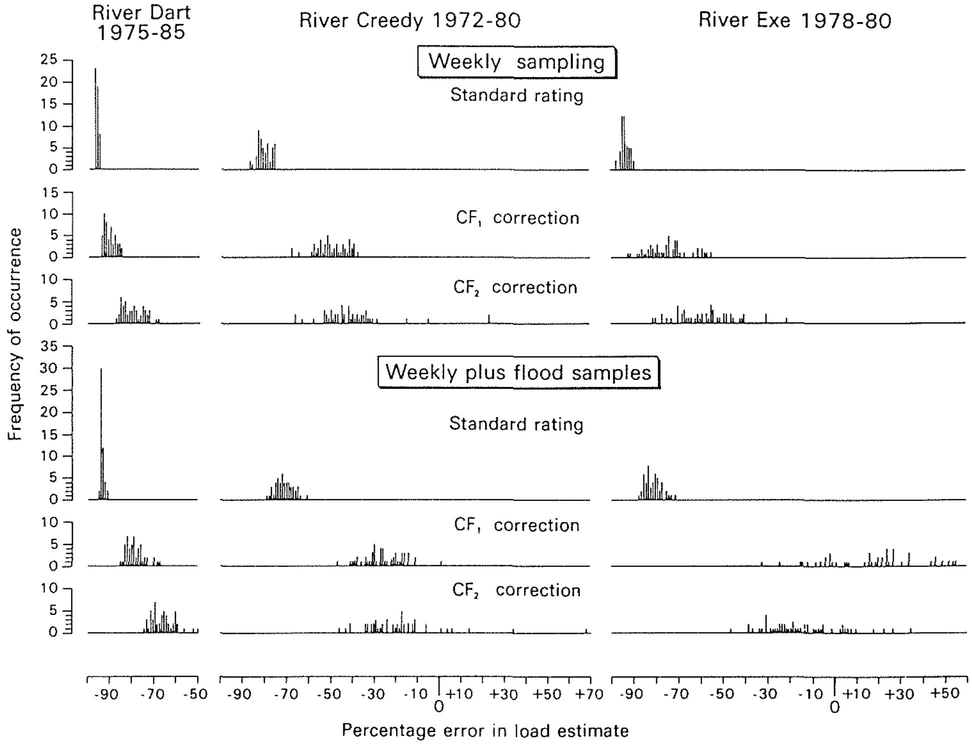

Fig. 1 The percentage error associated with the standard rating curve estimates and bias corrected estimates of suspended sediment load for the total period of record, obtained using rating curves fitted to the 50 replicate sets for both regular weekly sampling and regular weekly plus flood event sampling. The percentage error has been calculated as (estimated load/actual load – 1) × 100

estimates of annual and total suspended sediment yield for the three rivers investigated. In most cases, the estimates obtained significantly underestimated the actual loads, and the replicate results indicate an appreciable lack of precision. Furthermore, the considerable interannual variability in the level of accuracy of the load estimates obtained using the correction factors could provide significant problems in any attempt to analyse long-term trends in sediment records. For example, Tables 3 and 4 indicate that whereas the actual annual suspended sediment load of the River Creedy in water years 1974/1975 and 1977/1978 were almost identical, the rating curve estimates would be likely to indicate that the load for 1977/1978 was double that for 1974/1975. Figure 1 summarizes many of the findings presented in Tables 3 and 4 by plotting the percentage error of the 50 replicate load estimates obtained for the total period of record for each river using both standard and bias corrected rating relationships fitted to the data sets synthesized to represent regular weekly and regular weekly plus flood event sampling. These plots emphasize the lack of accuracy associated with most of the rating curve estimates, both corrected and uncorrected, and indicate that any improvement of accuracy is generally associated with a reduction in precision. They also provide further examples of instances where the bias correction procedure advocated by Ferguson (1986a, 1987) fails to provide reliable estimates of suspended sediment yield. In addition, the considerable variation in the accuracy of the rating curve estimates of sediment yield between the three rivers must be noted. Although some variation in the relative accuracy of the estimates between the different rivers seems inevitable in view of their different sized basins, a greater degree of consistency might have been expected for three adjacent basins experiencing similar hydrological conditions.

Ferguson (1986b) proposes that random sampling error and sampling bias may provide a cause of inaccuracy in rating curve load estimates. He suggests that in cases where the removal of the systematic error associated with transformation bias fails to yield accurate estimates, this failure may be due either to sampling error (i.e. imprecision in the rating curve estimates) or to the points defining the rating relationship not representing a random sample of all possible points. These assumptions may be readily tested by comparing the actual loads with estimates derived using rating curves established for the entire data set of concentration and discharge values for the period of record. Use of the entire data set removes both sampling error and sampling bias. The estimates of total load obtained for

Table 5 *Estimates of total suspended sediment load for the period of record obtained for the three rivers using rating relationships fitted to the entire data set of hourly discharge and concentration values*

River	*Period*	*Actual load (t)*	*Rating estimate (t)*	*Bias corrected (CF_1) rating estimate (t)*
Dart	*1975-1985*	*24 499*	*862*	*2 145*
Creedy	*1972-1980*	*82 863*	*15 443*	*39 579*
Exe	*1978-1980*	*41 402*	*2 754*	*9 212*

the three rivers using rating relationships developed from the entire data set of hourly discharge and concentration values are listed in Table 5. The estimates corrected for the systematic transformation bias using CF_1 range between 8.8% and 47.8% of the actual loads, again confirming that bias correction procedures are unlikely to provide accurate load estimates for these three rivers.

Figure 1 and Table 3, 4 and 5 suggest that bias associated with logarithmic transformation is not the prime cause of the inaccuracy of rating curve estimates of sediment yield for these rivers. Other factors not reflected in the correction factors are clearly important. These include the substantial scatter associated with the rating curve plots, seasonal differentiation of the rating relationships, lack of coincidence of the sediment concentration and discharge responses during storm events, and hysteretic and exhaustion effects (cf. Walling, 1977a, b). The influence of some of these factors is demonstrated in Fig. 2 for the River Dart. Figure 2A illustrates the considerable degree of scatter associated with a plot of the total data set of hourly concentration and discharge values for the period of record. A straight line fitted to this scatter plot explains only 14% of the total variance. Values of concentration associated with a given value of discharge range over more than three orders of magnitude, emphasizing the lack of any clearly defined functional relationship between concentration and discharge. Seasonal differentiation of the rating plot is also clearly evidence in Fig. 2A. Figure 2B demonstrates the lack of coincidence that may exist between the sediment concentration and discharge responses during flood events. During the first event, the peak of sediment concentration precedes the hydrograph peak. Use of the rating relationship, however, assumes that the two coincide and would in this instance tend to overestimate the load, if the values of concentration predicted by the rating curve were similar to the actual values. An exhaustion effect is also evident in Fig. 2B which shows a progressive decline in the sediment concentrations associated with the sequence of storm hydrographs. The complexity introduced into the rating relationship by such exhaustion effects and also by hysteresis effects is clearly shown in Fig. 2C which plots the trend of the rating relationship during the period illustrated in Fig. 2B. Hysteretic loops are clearly evident, emphasizing the inappropriateness of a simple straight line relationship. These loops show a progressive shift towards the right in response to exhaustion as the sequence of storm peaks proceeds, although there is some evidence of "recovery" between peaks 2 and 3.

Another cause of the inaccuracies associated with the use of rating relationships is the fact that a large proportion of the total suspended sediment load is transported by a few major flood events which represent only a very small proportion of the total time. Figure 3 presents plots of cumulative percentage sediment load volume *vs.* time, derived from the ranked series of load values, for the three rivers included in the previous analysis. These indicate that 75% of the total load is transported during approximately 1%, 2% and 4% of the time by the Rivers, Dart, Creedy and Exe respectively. These durations are equivalent to about 4, 7 and 14 days per year. This has important implications for the likely accuracy of rating curve estimates of suspended sediment load. Firstly it means that a regular sampling programme is unlikely to provide samples representative of those

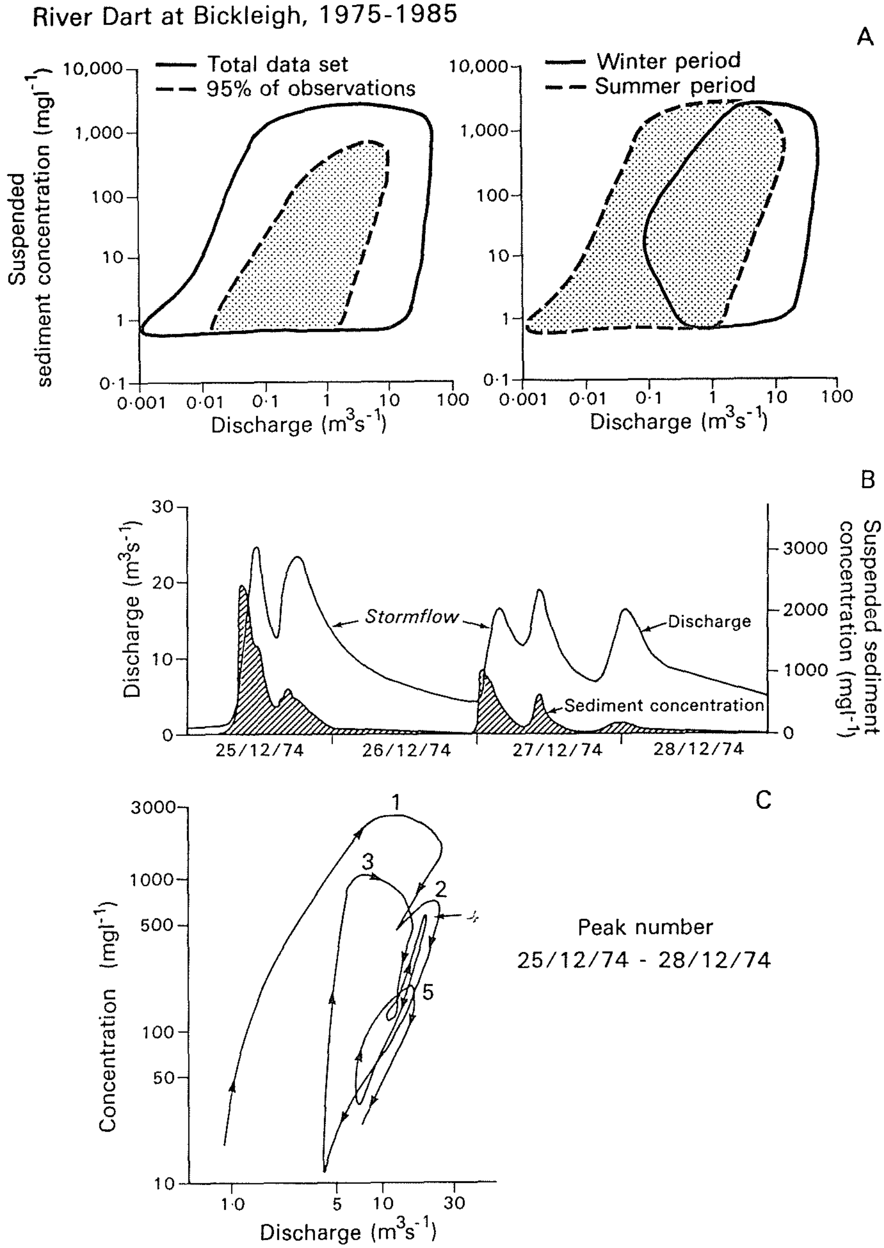

Fig. 2 The complex relationship between suspended sediment concentration and discharge exhibited by the River Dart at Bickleigh. A illustrates the scatter of the rating plot representing the entire data series for the period of record and the seasonal differentiation evidenced by the plot. B illustrates the lack of coincidence in timing that may exist between sediment and discharge response and the progressive exhaustion of sediment supply that may occur during a sequence of storm hydrographs. C plots the trend of the rating relationship during the period covered by (B).

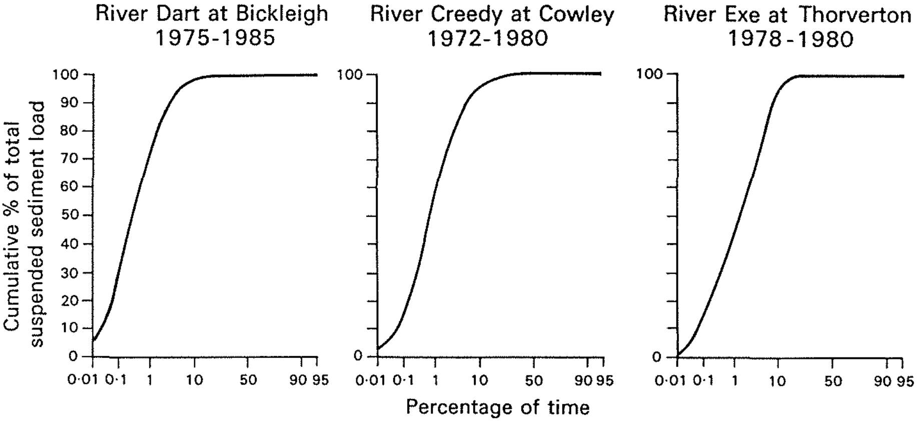

Fig. 3 Plots of cumulative percentage sediment load vs. time derived from the ranked series of load values for each of the three rivers.

periods when the majority of the load is transported. Secondly it means that because the rating plot is fitted by least squares to the whole range of discharge and concentration data, its trend may be largely determined by the main mass of samples representative of low discharges and concentrations, and may therefore be unrepresentative of the conditions during which the majority of the load is transported. The first implication is clearly evidenced by the results of the preceding analysis shown in Tables 3 and 4 and Fig. 1. The degree of underestimation was reduced when the data sets provided by regular weekly sampling were supplemented by samples collected specifically during flood events. However, even these data sets would have been biased towards low and medium values of concentration and discharge.

CONCLUSION

In an attempt to provide further evidence concerning the reliability of estimates of suspended sediment yield derived using rating curves, and more particularly the accuracy of load estimates corrected for transformation bias, this study has compared the actual sediment yields of three rivers in Devon, UK, with equivalent estimates obtained using rating curves. The results confirm the low levels of accuracy and precision reported for load estimates obtained for rating curves by the authors in a previous paper (Walling & Webb, 1981) and indicate that the use of bias correction procedures is unlikely to produce a major improvement in the reliability of the load estimates obtained for rivers experiencing similar conditions. Further attention could be given to other approaches to improving the accuracy of rating curve estimates, including the use of regression weighted towards the high values of concentration and discharge responsible for transporting the majority of the load. However, the substantial scatter evidenced by most rating relationships and the complexities associated with hysteresis and exhaustion effects are thought to preclude any major improvements in the reliability of rating curve estimates.

Accurate estimates of suspended sediment load are essential for sediment budget investigations, since it may be necessary to compare the sediment yields of sub-basins, to document the downstream changes in sediment load occasioned by conveyance losses, or to investigate temporal trends in sediment yield. The provision of reliable load estimates will in many cases necessitate the use of continuous turbidity monitoring equipment, as employed in this study, frequent sampling, or the use of sampling strategies specifically designed to produce accurate load estimators (cf. Thomas, 1985, 1986).

Acknowledgements Most of the sediment data used in this study were collected as part of investigations of the sediment dynamics of local rivers supported by the Natural Environment Research Council. The assistance of South West Water in providing river flow data for the rivers Exe and Creedy and of Mrs Sue Milward with computer programming is gratefully acknowledged.

REFERENCES

Adams, J. (1980) High sediment yield from major rivers of the western Southern Alps, New Zealand. *Nature, Lond.* **287**, 88-89.

Ashmore, P. E. (1986) Suspended sediment transport in the Saskatchewan River basin. *Sediment Survey Section, Water Survey of Canada Report IWD-HQ-WRB-SS-86-9.*

Colby, B. R. (1956) Relationship of sediment discharge to streamflow. *USGS Open File Report.*

Duan, N. (1983) Smearing estimate: a nonparametric retransformation method. *J. Am. Statist. Ass.* **78**, 605-610.

Farr, I. S. & Clarke, R. T. (1984) Reliability of suspended load estimates in chalk streams. *Arch. Hydrobiol.* **102**, 1-19.

Ferguson, R. I. (1984) Sediment of the Hunza River. In: *The International Karakoram Project* (ed. by K. J. Miller), vol. 2, 581-598. Cambridge University Press.

Ferguson, R. I. (1986a) River loads underestimated by rating curves. *Wat. Resour. Res.* **22**, 74-76.

Ferguson, R. I. (1986b) Reply. *Wat. Resour. Res.* **22**, 2123-2124.

Ferguson, R. I. (1987) Accuracy and precision of methods for estimating river loads. *Earth Surf. Processes and Landforms* **12**, 95-104.

Griffiths, G. A. (1979) High sediment yields from rivers of the western Southern Alps, New Zealand. *Nature, Lond.* **282**, 61-63.

Hansen, D. & Bray, D. (1987) Generation of annual suspended sediment loads for the Kennebacasis using sediment rating curves. In: *Proc. 8th Canadian Hydrotech. Conf, Can. Soc. Civ. Engrs* (Quebec, May 1987).

Kellerhals Engineering Services (1985) Sediment in the Pacific and Yukon region: review and assessment. *Water Resources Branch, Inland Waters Directorate, Environment Canada, Report IWD-HQ-WRB-SS-85-8.*

Koch, R. W. & Smillie, G. M. (1986) Comment on "River loads underestimated by rating curves" by R. I. Ferguson. *Wat. Resour. Res.* **22**, 2121-2122.

Olive, L. J., Rieger, W. A. & Burgess, J. S. (1980) Estimation of sediment yields in small catchments: a geomorphic guessing game? *Proc. Conf. Inst. Austral. Geographers* (Newcastle, New South Wales).

Stott, T. A., Ferguson, R. I., Johnson, R. C. & Newson, M. D. (1986) Sediment budgets in forested and unforested basins in upland Scotland. In: *Drainage Basin Sediment Delivery* (Proc. Albuquerque Symp., August 1986), 57-68. IAHS Publ. no. 159.

Thomas, R. B. (1985) Estimating total suspended sediment yield with probability sampling. *Wat. Resour. Res.* **21**, 1381-1388.

Thomas, R. B. (1986) Calibrating SALT: a sampling scheme to improve estimates of suspended sediment yield. In: *Monitoring to Detect Changes in Water Quality Series* (Proc. Budapest Symp., July 1986), 79-88. IAHS Publ. no. 157.

Walling, D. E. (1977a) Assessing the accuracy of suspended sediment rating curves for a small basin. *Wat. Resour. Res.* **13**, 531-538.

Walling, D. E. (1977b) Limitations of the rating curve technique for estimating suspended sediment loads, with particular reference to British rivers. In: *Erosion and Solid Matter Transport in Inland Waters* (Proc. Paris Symp., July 1977), 34-48. IAHS Publ. no. 122.

Walling, D. E. & Webb, B. W. (1981) The reliability of suspended sediment load data. In: *Erosion and Sediment Transport Measurement* (Proc. Florence Symp., June 1981), 177-194. IAHS Publ. no. 133.